

University of Leeds

Research Culture Awards 2021-22

Application Form

Please email nominations/self-nominations using this form to Holly Ingram H.L.Ingram@leeds.ac.uk by Wednesday June 1st, 5pm

Information section

This section covers information about the lead applicant and the team as well as the chosen category of award. Expand the sections as necessary. The maximum number of people in a team is set to 10 for event planning and personal prize budgeting purposes.

1. Select the nomination type

- Self-nomination
- ~~Third party nomination (you will need to secure the nominees' agreement for submission)~~
 - ~~○ Nominator name:~~
 - ~~○ Nominator School:~~
 - ~~○ Nominator role/post:~~

2. Lead applicant and team

Provide details about the applicant and the team (including external partners if applicable), and their contribution to the initiative. The lead applicant will be the contact person for the management of the award application.

Lead applicant name: Dr Emily Ennis

Lead applicant contributor role: Co-Manager of the Research Management Process in CDRC, responsible for overseeing outputs and impact developed from CDRC projects

Lead applicant Service or School / Faculty: Consumer Data Research Centre (CDRC)

Lead applicant role/post: Research and Impact Manager

Lead applicant career stage: ~~ECR, Mid-Career, Professor, Professional / Technical~~ staff

Team member name (include lead applicant)	Contributor role	School/Service/external organisation + role/post	If member deserves special mention, state reason (optional)
Professor Mark Birkin	Co-Director of CDRC (PI on ESRC grant)	Professor of Spatial Analysis and Policy, School of Geography; Director, Leeds Institute for Data Analytics; Co-Director, CDRC	
Professor Ed Manley	Co-Director of CDRC	Professor of Urban Analytics, School of Geography; Co-Director, CDRC	
Dr Nik Lomax	Co-Director of CDRC	Associate Professor of Data Analytics for Population Research, School of Geography; Co-Director, CDRC	

Adam Keeley	Oversaw creation of LASER (see below) and responsible for data governance	Data Analytics Team Manager; Formerly CDRC Data Scientist	
Dr Peter Baudains	Data scientist overseeing development of data products	Research Data Scientist, CDRC	
Kylie Norman	Lead co-ordinator on the data scientist development programme, overseeing capacity building and upskilling	Senior Operations Co-ordinator, CDRC Data Scientist Development Programme Co-ordinator, Leeds Institute for Data Analytics (LIDA)	
Robyn Naisbitt/ Mel Green	Lead for dissemination, public engagement, press outreach	Communications and Public Engagement Officer, CDRC	
Oli Mansell	Responsible for all data service requests, responsible for research management process in CDRC	Centre Manager, CDRC	
Paul Evans	Responsible for forging and sustaining relationships with external partners	Formerly Business Development Manager (CDRC), now Research and Innovation Development Manager (LIDA)	
Dr Emily Ennis	Research and Impact Manager	CDRC	

Expand table as necessary; the upper limit for a team is 10 members.

3. Select the award category that best fits the research culture activity:

- ~~Personal development, reward and recognition~~
- Open research and impact
- ~~Equality, diversity and inclusion in research~~
- ~~Responsible research and innovation~~
- ~~A collegiate and supportive environment~~

Case for Award

The Application form from this point on (= Case for Award) should not be longer than two pages in total (11pt Calibri).
Please write clearly, free from jargon, for a non-specialist reader.

1. Title

Opening up data science to solve real-world problems

2. Summary

CDRC is a research centre fostering collaboration with external partners through our ladder of engagement model, which has been called innovative by ESRC. As a data provider, we build trust and accountability through rigorous governance and infrastructures, including our virtual research environment, Leeds Analytics Secure Environment for Research (LASER), and our research management process. We encourage transparency and reproducibility by creating diverse types of derived data products, aimed at diverse groups, from policy makers and researchers, to activists and children. We publish on open platforms, including GitHub, in order to increase research reproducibility and to foster inclusivity beyond academia.

3. Why?

We provide the infrastructure for researchers wishing to access consumer data where a lack of partnerships with external data providers, funds required to purchase such data, or trusted research environment might be obstacles. As a data store, we work with over 30 data owners to make consumer data available to trusted researchers. We offer data under three tiers: Open, Safeguarded, and Secure. Access to Safeguarded/Secure data is through a reviewed application process. Secure data is accessed remotely by LASER or through our labs in Leeds, London, and Liverpool. The transition to a remotely-accessible virtual research environment – LASER – was accelerated by COVID, but has also expanded access to researchers unable to visit our safe rooms. This demonstrates a commitment to open research. By increasing accessibility and transparency to consumer data we offer greater opportunity for inclusive research practices, which is also evidenced in our data science training provision. We run online tutorials through the data store, options to participate in the Masters Dissertation Scheme, our [Consumer Analytics and Marketing Strategy MSc \(CAMS\)](#), our Data Analytics and Society [Centre for Doctoral Training](#), and the LIDA [Data Scientist Development Programme](#) (DSD, formerly known as the Internship Programme) which hosts CDRC projects. These are interdisciplinary courses, drawing on expertise from different faculties, which encourages collaboration across the University. We also run [CPD-accredited data science training](#) aimed at external stakeholders, making academic data and expertise available to other sectors. [CDRC's GitHub](#) also increases access to derived data products that can be used by researchers outside CDRC. Such upskilling and training provision have been developed in response to an identified lack of data science skills in the UK workforce, as highlighted in the [UK's National Data Strategy](#).

4. What?

CDRC's open research culture is underpinned by three key initiatives that helps expand access to high quality consumer data: collaboration/engagement, infrastructure, and outputs/impact. CDRC cannot exist without collaboration: it works jointly with UCL, Liverpool, and Oxford, as well as other ESRC-funded data centres to produce high quality research. It is also interdisciplinary, with Co-Investigators at the Centre spanning three faculties at Leeds – Environment, Medicine, and the Business School – and with ongoing research projects co-supervised in Faculties of Biological Sciences, Social Sciences, and Engineering and Physical Sciences. This promotes open and inclusive models of engagement within the University. With external organisations we also employ what we have named a 'ladder of engagement model' (see Q.5), which encourages external organisations to work with us, including depositing their data with us, on different-sized projects with varying and flexible levels of data sharing and collaboration. CDRC has a business development strategy that helps identify and target these partnerships outside academia, and the Partnerships Development Manager, as well as colleagues in RIS, provide expertise in brokering and managing these collaborations. The ladder of engagement model is effective as we guarantee high levels of data security for our partners, with data held and accessed securely, either in our safe rooms or through LASER, and overseen by CDRC's Research Data Scientist team and LIDA's Data Analytics Team. Research projects that use CDRC data are managed by the Centre Manager and Research and Impact (R&I) Manager to ensure research is carried out ethically. We prioritise making research insights available through open access or derived data products, and this is one of our key objectives in our current phase of funding (April 2022-September 2024). We also diversify the target demographics of our research by creating novel data products (such as [an online game](#) to help children understand the carbon footprint of their school dinners) and by using engagement opportunities such as the Be Curious and Leeds Data Festivals to showcase our research to the public. These strategies feed into UKRI research priorities, with ESRC's DigitalFootprints programme, beginning in 2024, focusing on how the economy and society respond to an increasingly-digital world.

5. How?

Cross-sectoral data sharing is difficult, with commercial sensitivities and competition law limiting how data can be shared and with whom. CDRC makes data sharing between the private sector and academics possible by offering responsible data governance and by promoting innovation. Researchers hoping to access the controlled data held by CDRC can only do so if they demonstrate that data will be used ethically and for the public good (as determined by our advisory board) and if they pass Safe Researcher Training (provided by UK Data Service or the Office for National Statistics (ONS)). Following this, their research project is entered into the CDRC's Research Management Process overseen by the Centre Manager and R&I Manager, with technical support provided by LIDA's Data Analytics Team. This level of security has enabled private sector data owners to commit to collaborative research and data sharing. The level of such collaboration and data sharing is determined by our Ladder of Engagement model: partners can work with CDRC at any "rung" of that ladder, beginning with very little data sharing (e.g. through a DSD Programme) with the prospect of moving along to a multilateral data sharing agreement (DSA). Private sector partners also value the insights generated by our research, and this determines the levels at which many collaborators choose to work with us. We therefore focus on expanding access to research findings. As an example, a research project on the Data Scientist Development Programme called Local Data Spaces (in collaboration with the Joint Biosecurity Centre (JBC), the ONS, ADR UK, and CDRC Liverpool) worked collaboratively with local authorities to solve problems relating to the COVID-19 pandemic. One of its outputs was a freely-accessible repository on GitHub containing code and data from the project, which meant that local authorities and councils could continue to draw insights after the project ended.

6. So what?

We produce quarterly and annual KPIs that provide quantitative insights into the number of: partnerships, derived data products, and projects active in LASER. CDRC hosts 30+ unique data sets and makes over 30 derived data products available to the public (with more anticipated by the end of 2022). Since 2016, we have also licenced safeguarded and controlled data to 23 internal projects and 29 external projects, with a total of 357 academic named on research projects and 4 partners from other sectors. We also measure our success through further funding received: over £188,014,305 for CDRC-related projects since February 2019. Our Research Management Process is designed to measure more qualitative outcomes, such as impact testimonials, which are captured by the R&I Manager.

7. What next?

CDRC has received further ESRC funding until September 2024 and our priorities will be to build reputation for our research outputs and impact. We will be using the Business Development Strategy, Ladder of Engagement model, and R&I Manager expertise to build further strategic partnerships and to strengthen and deepen existing relationships, as well as working to address national data research needs as part of the DigitalFootprints programme. Our Research Data Scientist team has recently expanded, increasing our resource to create derived data products. We are also re-committing to upskilling the data workforce through our programme of research-based education (see Q.3), having recently appointed a Teaching Fellow to consolidate and review our training following an evaluation from an external training provider and our successful CPD accreditation.

8. What challenges did you have in planning/organising/running/evaluating your initiative and how did you overcome these?

As an ESRC-funded data centre, our strategic priorities have been limited by our funding schedule. We now hold bi-monthly meetings with ESRC representatives and routinely share KPIs, in order to ensure CDRC objectives form part of their research priorities and funding allocation over the next 10 years. We have also built strong relationships with the RIS legal team to develop rigorous DSAs that have eased some challenges surrounding our partners' hesitancy to share data.

9. Was there something particularly innovative/creative about your initiative?

Because cross-sectoral data sharing can only be made possible with robust technical and legal infrastructure in place, CDRC research has depended significantly on professional and technical staff for its success. However, such dynamic and inclusive working patterns have enabled CDRC to sign master collaboration agreements with high profile UK retailers – the first of its kind at a UK university – and confidently engage partners who have previously been reluctant to work with universities. What makes the research culture open, transparent, inclusive, and efficient is how people and infrastructure work together to make it possible.

10. What makes your activity a notable example of culture change?

CDRC's approach to open research and impact has effected culture change in four ways. Firstly, by creating better mutual understanding of stakeholder requirements and approaches and the research process. Secondly, by focusing on real world problems (and data sources) at all levels in teaching and research. Thirdly, by producing high impact publications with non-academic authors. Finally, CDRC forges new and deep strategic partnerships at the top of the ladder of engagement, having recently signed a master collaboration agreement with a UK retail partner.